DOCUMENT RESUME

ED 205 121 HE 014 147

AUTHOR Schomberg, Steven F.; And Others
TITLE Using the College Outcome Measurement Project to Measure College Outcomes. AIR Forum 1981 Paper.
PUB DATE May 81
NOTE 21p.; Paper presented at the Annual Forum of the Association for Institutional Research (21st, Minneapolis, MN, May 17-20, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *College Seniors; Comparative Analysis; *Educational Benefits; Higher Education; Institutional Research; Majors (Students); *Outcomes of Education; *Research Methodology; Self Evaluation (Individuals); Student Motivation; *Undergraduate Study
IDENTIFIERS *AIR Forum; *College Outcome Measures Project; University of Minnesota

ABSTRACT

The usefulness of the College Outcome Measures Project (COMP), sponsored by the American College Testing Program, was assessed based on a 1979 study of 96 University of Minnesota graduating seniors. Three questions were examined: whether graduates within the same institution but in different curricula achieve different outcome scores; whether students from different institutions score differently on the COMP; and whether it is possible to substitute COMP for a locally developed self-report questionnaire on outcomes of undergraduate education. It was found that seniors in four different curricula (humanities, social sciences, natural sciences, and Inter-College Program) did not score differently. Possible reasons why COMP's previous field-test findings of differentiation among majors were disconfirmed are noted. COMP distinguished among groups that differed on level of academic achievement and motivation; COMP scores were related to self-reported academic achievement and aspiration. It is suggested that if the COMP test had been more difficult, the effects of grades on COMP scores would have been greater. COMP scores were found to be unrelated to student scores on self-perceived benefits of college and satisfaction with their skills and knowledge. The participants felt that the content of COMP was somewhat superficial and did not allow them to demonstrate their expertise. It is suggested that COMP has greater potential usefulness in distinguishing among institutions than in comparing outcomes of programs within an institution. Recommendations for conducting outcomes studies are offered. A bibliography is appended. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED205121

38

Using the College Outcome Measurement Project to Measure College Outcomes

Steven F. Schomberg, Darwin D. Hendel, and Caroline L. Bassett
University of Minnesota

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Association for Institutional Research
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Running Head: Using the College Outcome Measurement Project

HE 014 147

THE ASSOCIATION FOR INSTITUTIONAL RESEARCH

This paper was presented at the Twenty-First Annual Forum of the Association for Institutional Research held at the Leamington Hotel in Minneapolis, Minnesota, May 17-20, 1981. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Mary Corcoran
University of Minnesota
(Editor, AIR Forum Publications)

27 A(2)

Abstract

This paper describes one institution's use of instruments developed by the College Outcome Measures Project (COMP), a project sponsored by the American College Testing Program. The study was designed to answer three questions concerning the use of COMP. Results obtained in the study of 96 University of Minnesota graduating seniors in the 1979 national field test of COMP suggest greater potential usefulness in distinguishing among institutions than in comparing outcomes of programs within an institution. The authors conclude that as new approaches and instruments for measuring college outcomes become available to institutional researchers, criteria for instrument selection need to be shared among individuals charged with conducting outcome studies.

Using the College Outcome Measures Project To Measure College Outcomes

The general belief among educators and institutional researchers is that an undergraduate education does have an impact on students. That belief is supported by data, such as that contained in Feldman and Newcomb's (1973) review of outcome studies conducted from the 1920s to the 1960s. They concluded that colleges do have an impact on students' personality, attitudes and values, and that the consequences often persist after the college years. In a more recent review, Bowen (1978) examined college outcomes within the context of the cost of higher education. Bowen argued that outcomes of higher education consist primarily of changes that inevitably influence the character of society. Not only do college graduates earn more than non-graduates, but they also add to society's human capital with their acquired skills and abilities, without which our society could not make its social, economic or cultural advancement.

While outcome studies are useful for stimulating general discussion of liberal education, educational researchers often wish to assess the specific effects of their institution's curriculum. Astin's (1977) student-outcomes taxonomy is useful in distinguishing among the various assessment procedures which could be used. According to Astin, decisions must be made about (a) type of outcome (cognitive or non-cognitive), (b) type of data (psychological or behavioral), and (c) the length of time of the effect of college on the outcomes (short-term versus long-term). A follow-up study of graduates has been the typical approach

used in assessing the specific effects of an institution's curriculum.

Colleges and universities are giving increasing attention to the outcomes of undergraduate education, especially to the generic outcomes of a general or liberal education. This concern for liberal education outcomes has occurred in response to at least three trends. First, today's students are turning away from bachelor's degree in letters and arts, social sciences, and foreign languages, and toward degrees in specific career and professional areas such as computer and information services, health professions, business, and engineering (Roemer, 1981). Second, critics argue that the elective system has diluted the concept of a liberal education beyond meaningful definition (Hechinger, 1978). Third, the demand by the public for efficiency and accountability has made institutions of higher learning cost- and purpose-conscious.

Two approaches for measuring college outcomes can be found in the literature: a) the use of assessment instruments developed and normed by professional test development organizations; and b) the use of instruments developed locally and tailored for a specific institution. The present study addresses the relative usefulness of the two approaches for measuring college outcomes. The College Outcome Measures Project (COMP), developed by the American College Testing Program (ACT) with initial assistance from the Fund for the Improvement of Post-Secondary Education, is a psychologically oriented test that assesses long-term cognitive outcomes. The COMP is intended to measure knowledge and skills appropriate to (a) measuring the impact of a liberal education, and (b) defining what constitutes an effectively functioning adult in society. ACT developed COMP by assembling a panel of educators from nine institutions to advise them concerning outcome dimensions and item content.[1]

Results of field studies of COMP (ACT, 1977; ACT, 1978; ACT, 1981) indicate that COMP differentiates between groups of college seniors and matched groups of vocational-technical students, between freshmen and seniors, and among majors. COMP is promising as an assessment procedure for measuring college outcomes, although published studies using COMP are just now appearing. At the Tennessee Technological University, Dumont and Troelstoup (1980) correlated student self-reported general education outcomes with measured outcomes on COMP. They found that among the 93 seniors who completed the two types of outcome measures, the highest correlation between scores from similar content areas was .34. Their findings imply a lack of validity of either self-report measures or of COMP, and suggest that institutions wishing to study general or liberal education outcomes use both types of measures.

The present study was designed to examine three questions: (1) Do graduates within the same institution but in different curricula achieve different outcome scores? (2) Will students from different institutions score differently on the COMP? (3) Is is possible to substitute COMP for a locally developed self-report questionnaire on outcomes of undergraduate education?

Method

Sample. Volunteers for the study consisted of 96 University of Minnesota seniors who had completed 160 or more quarter credits (of 180 credits needed for graduation) and were enrolled in the College of Liberal Arts (CLA), the General College (GC), or the Inter-College Program (ICP). The CLA students were enrolled in a regular major curriculum in one of three general subject matter areas: humanities, social sciences, or natural sciences.

Students in the individually designed programs in GC and ICP often include career-related educational experiences in their curricula.

Based on past experiences in soliciting student volunteers to complete tests, we expected that about one-third of the students invited would participate. Because we wanted 30 students or more from each group for data analysis, we invited 90 students to participate in the testing. Where the number of seniors meeting the selection criterion was 90 students or fewer, all students were invited to participate. Table 1 shows the sampling plan and the actual number of participants. Comparisons between the volunteers and non-volunteers on their cumulative grade-point average yielded no statistically significant ($p > .05$) differences in grade-point average between the two groups (volunteers -- N = 93, $\bar{x}$ = 2.98, SD = .51; non-volunteers -- N = 1,336, $\bar{x}$ = 2.94, SD = .92).

Insert Table 1 About Here

Instrumentation. During the spring of 1979, students completed the Objective Test, one of the assessment instruments that constitute the COMP system. The items of the Objective Test represent six domains: three are content-specific (Functioning with Social Institutions, Using Science and Technology, Using the Arts) and three are process-specific (Communicating, Solving Problems, Clarifying Values). Unique simulation activities and interpretative exercises test students' abilities to function as effective adults in dealing with a variety of issues and problems. For example, one item asks the student to act as a member of a library board and to recommend one of two art prints to the board for acquisition by

the library. The two-hour test includes 63 items, each with four multiple-choice answers, two of which are correct for each item, one in a content domain and the other in a process domain; scoring for each item ranges from -2 to +2.

Validity data for the Objective Test were collected as part of the field test of which this study is a part and has just been reported by ACT (1980). ACT bases its validity argument for the Objective Test on the extent to which the Objective Test correlates with the Composite Examination. The Composite Examination is similar to the Objective Test, but is longer and contains open-ended answers, both oral and written, as well as multiple-choice answers. ACT believes that the Objective Test serves as a proxy measure for the Composite Examination and reports correlation coefficients between the two tests ranging from .64 to .88. Data collected from 2,265 students who completed the Composite Examination indicate near zero correlations with age (r = .10), weak (r = .21) correlations with cumulative grade-point averages, and moderate to high correlations with ACT Composite scores (r = .66 for 802 freshmen and r = .47 for 502 seniors). Finally, seniors in the three general COMP content areas (humanities, social sciences, natural sciences) had scored higher in their respective areas of interest than in the other two.

Since each item in the Objective Test is scored for both content and process domains, split-half reliability estimates are inappropriate. Equivalent-forms reliability estimates are .79 for the total scores and from .56 to .78 for the six subscores (ACT, 1981). Median intercorrelations among the six subscores are .69 for a group of 933 freshmen and .66 for a group of 915 seniors, although the two-dimensional nature of the scoring

procedures makes difficult a straightforward interpretation of the subscore intercorrelations. No stability data have been reported.

Participants in the present study also completed a 19-item perceived benefits questionnaire (e.g., "broadened acquaintance with improtant literature") adapted from the "Former Student Survey" used at the University of Minnesota. Participants provided information on their gender, major, grades in college, educational aspiration, satisfaction with the University, and satisfaction with their gains in skills measured by the COMP.

Procedures. Letters signed by their respective dean or director invited seniors to participate in the study. As an incentive, students were offered a one-dollar coupon to redeem at a local ice cream parlor upon completion of the test. Two issues of COMP NOTES were mailed to students to provide background on the COMP and to create in students the feeling that they were participating in a special study of national scope. Testing occurred in one of seven 2½ hour sessions of the students' choice during the third week of the 1979 Spring Quarter.

Data Analysis. The Objective Test was machine-scored by ACT and descriptive statistics (frequencies, percentages, means, and standard deviations) were computed for each subscore and for the total score. Comparisons among curricular and other groups were performed by using the analysis of variance program contained in the Statistical Package for the Social Sciences (SPSS). The GC students were excluded from the curricular comparisons because only five seniors completed the Objective Test. Correlational analysis was performed among scores on the items from the Former Student Survey and scores on the COMP tests. Also, correlations were calculated between COMP scores and data on students' satisfaction with

their undergraduate education and with their knowledge and skill level. For the purpose of describing scores, ACT compiled into percentile tables the scores of 915 seniors from the 15 institutions participating in the national field test.

Results

Differences among University of Minnesota seniors in four curricular groups were examined by looking at COMP total score and subscore means and standard deviations. Table 2 shows that seniors in four different curricula (humanities, social sciences, natural sciences, and Inter-College Program) did not have statistically significant differences among mean scores ($p > .05$). In fact, mean total scores for the four curricular groups were virtually identical, ranging from a low of 204 to a high of 208. Moreover, the subscore means did not show the expected differences in content subscore means among the curricular groups. The CLA natural science group had the lowest mean score on the Using Science subscore, and the CLA humanities and CLA social science groups had only slightly higher means for the Using the Arts and Functioning in Social Institutions subscores.

Insert Table 2 About Here

The 96 seniors were grouped on other dimensions that might relate to the content of undergraduate educational experiences. Comparisons yielded no statistically significant ($p > .05$) differences in COMP mean

scores between students who entered as freshmen and students who transferred from other institutions; between three age groups (20-22, N = 22; 23-30, N = 68; and 31 and over, N = 15); or between males and females.

Finally, we examined COMP scores in relation to indices of academic achievement and motivation: grade-point average and educational aspiration. We found statistically significant ($p \leq .05$) differences among three self-reported grade groups [mostly A's (N = 30), mostly B's (N = 36), and mostly C's (N = 17)] for the COMP total score ($F = 7.40$) and for each of the subscores. For the total score and for each subscore, the "mostly A's" group had the highest mean ($\bar{x} = 213.5$ for the total score), followed by the "mostly B's" group, ($\bar{x} = 202.5$ for the total score) and the "mostly C's" group ($\bar{x} = 198.4$ for the total score). Comparisons among educational aspiration groups yielded statistically significant ($p \leq .05$) differences for four of the subscores as well as for the total score. The trend was the same for all: mean scores were lowest for students who planned to receive a bachelor's degree, and highest for those who planned to pursue the doctorate.

The mean scores for all 96 University of Minnesota seniors were plotted on a norm table provided by ACT to compare University seniors with the 915 seniors from the 15 institutions that participated in the field study.[2] Means for University of Minnesota students were approximately one-half of a standard deviation above the overall mean for the COMP total score and for each of the six subscores. The COMP total score for University of Minnesota seniors was second highest of all fifteen institutions. The University of Minnesota seniors had the highest means for Using the Arts and Clarifying Values, the second highest for Using Science

and Solving Problems, the third highest for the Communicating scale, and the fourth highest for Functioning in Social Institutions.

We were interested in the relationship between COMP Objective Test scores and other locally relevant data concerning graduates' satisfaction with their undergraduate education. The first set of six satisfaction questions concerned students' satisfaction with their skills and knowledge in particular areas covered by the COMP Objective Test (e.g., being able to appreciate the arts). Although results indicate that students were "satisfied" with their skills and knowledge in most areas, there was considerable variability in their responses. The lowest mean (4.45) was for "being able to use science and technology" and the highest mean (5.21) was for "solving problems" (on a seven-point scale ranging from extremely dissatisfied to extremely satisfied). Comparisons among mean satisfaction scores for the four curricular groups indicated no statistically significant differences in satisfaction scores for any area except "being able to use science and technology" for which the CLA natural science group scored the highest.

We expected moderate, statistically significant correlations between students' self-reported satisfaction with their skills and knowledge in areas covered by COMP with their performance in parallel areas. The correlations suggest few relationships between the two sets of variables. Only four of the 42 correlations were statistically significant at $p \leq .05$, and the magnitude of the four correlations was small. The highest correlation ($r = .29$) was found between students' self-reported satisfaction in being able to appreciate the arts and their COMP subscore Using the Arts. The correlation between overall satisfaction with undergraduate

education and COMP was statistically significant ($p \leq .05$) for Using the Arts ($r = .25$) and Communicating ($r = .20$).

The final analysis correlated COMP scores with students' responses to a standard set of 19 benefit items (e.g., "broadened acquaintance with important literature") included on surveys of recent University graduates. We expected to find positive correlations between the two sets of variables. We found that a large majority (106 of 133) of the correlations were small and negative; the largest was $r = -.24$. These results suggest that there is a slight tendency for the students to perform less well on the COMP when they feel that their University undergraduate experiences have been beneficial to them than students who are less positive in their evaluations of their University undergraduate experiences.

Discussion

Results of the present study did not confirm COMP's previous field-test findings, which showed differentiation among majors if class was held constant; seniors had similar COMP scores regardless of their major or program. There are at least three possible reasons for this lack of differentiation. First, the test may have been too easy for many of the students, so that the test did not differentiate among students grouped toward the top of the range. Second, since all students had fulfilled distribution requirements--curricular experiences intended to effect general education outcomes--there may be little qualitative difference among students. Third, the way in which ACT conceptualized and measured some of the outcome areas (e.g., Using Science) may have been biased in

favor of certain beliefs (e.g., the importance of political and social aspects of using science and its technology rather than scientific facts.) COMP appears, as of now, not sophisticated enough for intra-institutional comparisons, but may prove useful in determining gross differences in the quality of general/liberal education between institutions. Additional inter-institutional studies must be conducted to substantiate the usefulness of COMP in determining overall institutional quality and in distinguishing among institutions and types of institutions.

Although COMP did not differentiate among University curricular groups, it did distinguish among groups that differed on level of academic achievement and motivation. ACT (1981) indicates that COMP measures college abilities not reflected by grade-point average; the assertion is based on a low correlation of COMP scores with grades (r = .21). In contrast, however, we found that COMP scores were related to self-reported academic achievement and aspiration. Furthermore, we believe that our data understate the case due to the very high scores achieved by most of the seniors. Had the COMP test been more difficult, we assume that the effects of grades on COMP scores would have been even greater. The inconsistency of our findings with ACT's suggests the necessity of further studies on the relationship of COMP to academic achievement and to prior measures of academic ability.

The COMP test alone does not appear to be a wholly adequate measure of outcomes specific to a particular college or university. We found COMP scores unrelated to University seniors' scores on satisfaction with their skills and knowledge in areas covered by COMP and with perceived benefits from their University education. In both cases the correlations

were small, and in the latter case the correlations showed a slight tendency to be negative. Our findings are consistent with those of Dumont and Troelstrup (1980) who reported that Tennessee Technological University student scores on a self-report measure of perceived benefits were independent of COMP scores. Researchers who wish to measure specific institutional outcomes might want to use locally oriented outcome measures along with COMP in order to assess outcomes not addressed by COMP.

A second concern relates to the acceptability to students and faculty of a particular assessment procedure. The 96 seniors in the present study felt that the content of COMP was somewhat superficial and did not allow them to demonstrate their expertise. A small number of faculty who participated in a related study concurred with students in their criticism of COMP and further argued that the six outcomes assessed by COMP did not put enough emphasis on critical thinking and broad familiarity with a variety of disciplinary areas. An institution considering use of COMP should give careful attention to the congruence between its ideas of the purposes of its general or liberal education curriculum and the outcomes measured by the COMP system.

As the pressures increase for demonstrating the impact of undergraduate education, other test development professionals will be providing a greater choice among standardized college outcome instruments. Given the diversity of higher education institutions across the nation, the development of a broad-based outcome measure, standardized on a national basis, that can also be used for intra-institutional evaluation, is a very difficult task. ACT's development of the College Outcomes Measures Project tests is a major step forward in providing such a measure of general and liberal education outcomes.

Footnotes

[1]Panel members came from Alverno College, Brigham Young University, Delaware County Community College, Florida International University, the State University System of Florida, Mars Hill College, Our Lady of the Lake University, the Tennessee Higher Education Commission, and the University of Nebraska-Lincoln.

[2]The norm group includes students from the following institutions: Baldwin-Wallace College, Drake University, Eastern Kentucky University, Maharishi International University, Michigan State University, Mississippi State University, Oklahoma City University, Rochester Institute of Technology, State University of New York-Geneseo, University of California-Los Angeles, University of Illinois-Urbana, University of Minnesota, University of Missouri-Columbia, University of Mid-America, and Valley City State College.

References

American College Testing Program. College outcome measures project annual report 1977. Iowa City, Iowa: ACT, 1977.

American College Testing Program. College outcome measures project annual report 1978. Iowa City, Iowa: ACT, 1978.

American College Testing Program. College outcome measures project revised summary report of research and development 1976-1980. Iowa City, Iowa: ACT, To be published, 1981.

Astin, A. W. Four critical years. San Francisco: Jossey-Bass, 1977.

Bowen, H. R. Investment in learning. San Francisco: Jossey-Bass, 1978.

Dumont, T. G. and Troelstrup, R. L. Exploring relationships between objective and subjective measures of instructional outcomes. Research in Higher Education. 1980, 12(1), 37-51.

Feldman, K. A. and Newcomb, T. M. The impact of college on students. San Francisco: Jossey-Bass, 1973.

Hechinger, F. M. Retreat from anarchy. Saturday Review, April 1, 1978, 18-19.

Roemer, R. E. Vocationalism in higher education: explanations from social theory. The Review of Higher Education. Winter 1981, 4(2), 23-46.

Table 1

Description of study design and sampling plan, by group of students

Curricular group	Total seniors	Seniors invited	Participants (N = 96)
CLA Humanities	352	89	19 (21.4%)
CLA Social Science	839	90	27 (30.0%)
CLA Natural Science	90	90	24 (26.7%)
General College	38	38	5 (13.2%)
Inter-College Program	109	91	21 (23.1%)

Note. Seniors were defined as having completed 160 or more credits.

Table 2

Means and standard deviations for COMP total score and subscores, total group of seniors and by group

COMP Score	Highest Possible Score	Total Seniors (N=96)		Groups[a]								F-ratio	p-value
				CLA Humanities (N=19)		CLA Social Science (N=27)		CLA Natural Science (N=24)		ICP (N=21)			
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD		
Functioning in Social Institutions	84	66.43	5.95	66.79	6.06	67.37	5.34	66.25	6.97	65.81	4.93	.32	.81
Using Science	84	70.27	6.44	69.58	6.36	72.19	6.02	69.42	6.81	70.43	5.84	1.03	.38
Using the Arts	84	68.77	6.53	69.58	6.51	68.67	6.62	68.50	7.46	69.38	5.63	.14	.94
Communicating	72	55.03	5.39	54.42	4.98	55.85	6.57	54.75	5.77	55.00	3.99	.30	.83
Solving Problems	96	81.03	6.83	82.11	6.11	82.33	5.84	79.54	8.23	81.24	5.96	.88	.46
Clarifying Values	72	61.14	5.54	61.84	4.69	61.63	4.87	61.08	6.97	61.10	5.32	.10	.96
TOTAL	252	205.47	15.73	205.95	14.78	208.22	15.39	204.17	18.77	205.62	12.58	.29	.83

[a]Because of the small size of the General College sample, it was not included in the one-way ANOVA among groups of students